Popular Summary of

"Evaluating the credibility of transport processes in the Global Modeling Initiative 3D model simulations of ozone recovery"

To be submitted to the Journal of Geophysical Research - Atmospheres

Authors: Susan E. Strahan
Code 916/Goddard Earth Science & Technology Center

Anne R. Douglass
Code 916/NASA GSFC

The Global Modeling Initiative has integrated two 35-year simulations of an ozone recovery scenario with an offline chemistry and transport model using two different meteorological inputs. Physically based diagnostics, derived from satellite and aircraft data sets, are described and then used to evaluate the realism of temperature and transport processes in the simulations. Processes evaluated include barrier formation in the subtropics and polar regions, and extratropical wave-driven transport. Some diagnostics are especially relevant to simulation of lower stratospheric ozone, but most are applicable to any stratospheric simulation.

The temperature evaluation, which is relevant to gas phase chemical reactions, showed that both sets of meteorological fields have near climatological values at all latitudes and seasons at 30 hPa and below. Both simulations showed weakness in upper stratospheric wave driving. The simulation using input from a general circulation model (GMI_{GCM}) showed a very good residual circulation in the tropics and northern hemisphere. The simulation with input from a data assimilation system (GMI_{DAS}) performed better in the midlatitudes than at high latitudes. Neither simulation forms a realistic barrier at the vortex edge, leading to uncertainty in the fate of ozone-depleted vortex air. Overall, tracer transport in the offline GMI_{GCM} has greater fidelity throughout the stratosphere than the GMI_{DAS}.

Evaluating the credibility of transport processes in the Global Modeling Initiative 3D model simulations of ozone recovery

Susan E. Strahan
Goddard Earth Science & Technology Center
University of Maryland, Baltimore County
Baltimore, MD 21250

Anne R. Douglass
Code 916
NASA Goddard Space Flight Center
Greenbelt, MD 20771

Abstract. The Global Modeling Initiative has integrated two 35-year simulations of an ozone recovery scenario with an offline chemistry and transport model using two different meteorological inputs. Physically based diagnostics, derived from satellite and aircraft data sets, are described and then used to evaluate the realism of temperature and transport processes in the simulations. Processes evaluated include barrier formation in the subtropics and polar regions, and extratropical wave-driven transport. Some diagnostics are especially relevant to simulation of lower stratospheric ozone, but most are applicable to any stratospheric simulation.

The temperature evaluation, which is relevant to gas phase chemical reactions, showed that both sets of meteorological fields have near climatological values at all latitudes and seasons at 30 hPa and below. Both simulations showed weakness in upper stratospheric wave driving. The simulation using input from a general circulation model (GMI_{GCM}) showed a very good residual circulation in the tropics and northern hemisphere. The simulation with input from a data assimilation system (GMI_{DAS}) performed better in the midlatitudes than at high latitudes. Neither simulation forms a realistic barrier at the vortex edge, leading to uncertainty in the fate of ozone-depleted vortex air. Overall, tracer transport in the offline GMI_{GCM} has greater fidelity throughout the stratosphere than the GMI_{DAS}.

1. Introduction

For the past few decades, chemistry and transport models have been used to assess the impact of natural and anthropogenic perturbations such as aircraft emissions or chlorofluorocarbon growth on stratospheric ozone. Most assessments relied on two-dimensional (zonally averaged) models that cannot physically represent inherently 3D processes such as transport out of the polar vortex and cross tropopause transport [*Park et al*, 1999]. Some recent efforts have used three-dimensional chemistry and transport models (CTMs), which provide more realistic representations of non-zonal processes [*Danilin et al.*, 1998; Douglass et al., 1999; *Kinnison et al.*, 2001], although the third dimension greatly increases the computational requirements and demands greater human resources for evaluation. In spite of this, development of a 3D assessment model is a worthwhile goal because it offers the opportunity to improve the physical basis of assessment modeling and, if the 3D model compares well against observations, reduce uncertainties due to transport.

The Global Modeling Initiative (GMI) was formed in 1995 with the goal of producing a well-tested 3D chemistry and transport model that could be used for assessments and other controlled experiments that required a common framework. In their first experiment, the GMI science team used the same offline chemistry and transport model [*Rotman et al.*, 2001] with 3 different sets of meteorological input to evaluate which input would provide the most realistic simulation of an emissions scenario [*Kinnison et al.*, 2001]. To establish which simulation was the most credible, physically based tests, derived from aircraft and satellite data sets, were used to evaluate the simulations [*Douglass et al.*, 1999]. Six tests were created, each evaluating a different aspect of stratospheric transport and mixing, and grading standards were defined by the observations and their uncertainties. The simulations received scores on each test that could then

be used to quantitatively distinguish between them. At the end of the evaluation, the GMI science team could objectively select the best meteorological data set for simulating the effects of supersonic aircraft on the stratosphere.

Recently, the GMI science team ran two 35-year integrations of the WMO scenario MA2 [*WMO*, 2002] with the GMI-CTM. During the period simulated, 1995-2030, the scenario's organic chlorine and bromine boundary conditions decline while N_2O and CH_4 increase. The GMI model, input meteorological fields, and the scenario simulated are described in detail in *Connell and Douglass* [2003]. While the intent of this WMO scenario is to predict future ozone change, the primary purpose of this GMI study is to assess the sensitivity of model predictions to differences in transport. We chose two inputs for these simulations, one from the Finite Volume General Circulation Model (FVGCM) and the other from the Finite Volume Data Assimilation System (FVDAS). The CTM calculations used 1-year of meteorological fields from each model, repeating them for the 35-year simulation; the CTM simulations will be referred to as GMI_{GCM} and GMI_{DAS}. These data sets were selected because although they have significant differences in residual circulation and mixing [*Schoeberl et al.*, 2003], each is also known to realistically represent some aspects of the stratosphere. Initial evaluations using the GMI grading criteria on Goddard CTM simulations using the FVGCM and FVDAS winds showed that both data sets have stratospheric transport characteristics superior to the previously tested GMI simulations.

In this paper we build on the observationally based model evaluation philosophy discussed in *Douglass et al.* [1999]. The emphasis here is not just 'comparison with data', but 1) identification of atmospheric processes relevant to the realistic simulation of a phenomenon or feature, and 2) identification of a data set that demonstrates the occurrence of this process. Such a data set becomes the physical basis for a model evaluation test. A model earns credibility when

it can be shown to realistically represent known atmospheric processes. Some of the atmospheric processes relevant to the WMO MA2 scenario can be tested using previously defined GMI evaluation criteria. In this paper, new physically based stratospheric diagnostics are presented that illustrate additional stratospheric processes. While some of the tests are especially relevant to this scenario with declining halogens, all tests are generally applicable to any stratospheric simulation. In the following sections, we identify some important stratospheric processes and their diagnostic tests, and apply them to two GMI 3D-CTM simulations in order to evaluate many (though not all) aspects affecting their credibility in an ozone recovery scenario. Applying these tests to simulations run in a common framework (i.e., the GMI-CTM) allows us to examine the sensitivity of the results to the meteorological input.

2. Evaluating the suitability of GMI simulations for use in ozone predictions

How do you evaluate a model's credibility? The GCM-Reality Intercomparison Project for SPARC (GRIPS) compared temperature and wind fields in 13 3D middle atmosphere climate models against observations to identify deficiencies in dominant atmospheric features, such as the location of the jets and polar temperatures [*Pawson et al.*, 2000]. To evaluate an assessment model, one might also look at qualitative agreement with historical ozone trends. The Scientific Assessment of Ozone Depletion [*WMO*, 2002] shows many models' simulations of column ozone. While many models show qualitative agreement with historical ozone from 1980 to 2000, their predictions of future ozone diverge. Such agreement is a misleading diagnostic since the total column represents the integrated effects of chemistry and transport at many altitudes. To understand the difference in performance between 8 chemistry-climate models, *Austin et al.* [2003] chose several specific diagnostics, such as ozone climatology, polar temperature biases,

and poleward heat fluxes. Their intent, commensurate to that of GMI, is to find a range of diagnostics relevant to processes influencing ozone and use them to reduce the uncertainty in predictions of future ozone levels.

In this study we assess model credibility by evaluating temperature and transport processes. This is done with observationally based tests at a variety of altitudes and latitudes. In GMI-1, we developed tests to assess model transport processes in regions that would be perturbed by stratospheric aircraft exhaust. Since emissions were projected to occur in the upper troposphere and lower stratosphere, transport near the tropopause was especially important and tests were developed that emphasized model fidelity there. In this ozone recovery scenario, transport fidelity is important at all levels where Cl chemistry changes. For example, in the lower stratosphere, we expect changes in PSC processes to affect O_3, and in the upper stratosphere, we expect ozone loss by gas phase Cl reactions to be important. In this section, new tests are presented that expand the scope of the GMI's stratospheric evaluation. These tests, as well as some of the previous GMI tests, are applied to two new GMI-CTM simulations.

Data sets used in these tests are from the National Center for Environmental Prediction (NCEP)/National Center for Atmospheric Research (NCAR) Reanalysis, the Upper Atmosphere Research Satellite (UARS), and an ER-2 airborne spectrometer. NCEP/NCAR temperature analyses from 1980-1999 are used to create a climatology of monthly temperature distributions for 8 stratospheric levels and 11 latitude bands. (See *Newman et al.* [2001] for details of the reanalysis products.) UARS data sets include N_2O and CH_4 from the Cryogenic Limb Array Etalon Spectrometer (CLAES) [*Roche et al.*, 1996] and CH_4 from the Halogen Occultation Experiment (HALOE) [*Park et al.*, 1996]. Both instruments began operation October, 1991; CLAES made measurements for about 18 months and HALOE has operated nearly continuously

UARS was launched. CLAES has high spatial density sampling that alternates between 35°S-80°N and 35°N-80°S every 35 days. HALOE collects ~30 profiles daily with latitudes ranging from 80°S-80°N; latitudes poleward of 60° are only sampled in spring and summer. HALOE data sets are especially useful for investigating interannual variability and lower stratospheric transport.

2.1 Generalized tests of stratospheric temperature and transport

In this section, observationally based tests derived from aircraft and satellite observations are used to probe basic aspects of the stratosphere, such as temperature, transport, and mixing characteristics. The tests are used to identify strengths and weaknesses in the GMI-CTM simulations, but can be sensibly applied to any 3D online or offline simulation. In Section 2.2, we present tests that are particularly relevant to assessing the effects of the polar ozone loss.

2.1.1 Temperature

Both gas and heterogeneous phase chemical reactions are important in the stratosphere. Although both reaction types depend on temperature, they require different tests to evaluate model temperature behavior. Gas phase temperature-dependent reactions will proceed at a slightly slower or faster rate as temperature varies, but heterogeneous reactions only occur if the necessary temperature threshold is reached. (For regions where heterogeneous reactions are important, temperature distribution matters more than the mean. This evaluation will be discussed in Section 2.2.1.) To judge model temperatures for gas phase chemical reactions, we compare how often and how closely model temperatures agree with climatological values. This is accomplished by comparing the model and observed most probable temperatures for a given month, latitude, and altitude. Other evaluations of climate models have focused on polar temperatures or have looked at broadly averaged (*i.e.*, monthly, annual, or global) temperatures

[*Austin et al.*, 2003; *Pawson et al.*, 2000]. The temperature diagnostic shown here attempts to make a spatially and temporally thorough comparison using a little averaging as possible, whose results can be displayed as simply as possible. Model temperatures in the GMI simulations are a property of the input meteorological fields (i.e., of the FVGCM and the FVDAS).

Stratospheric temperature evaluation is based on the climatological mean distribution of NCEP/NCAR reanalysis temperatures from 1980-1999. For the entire 20-year data set, daily area-weighted temperature distributions are calculated for 11 latitude bands and 8 pressure levels from 150-1 hPa, and a mean distribution for each month is calculated. Climatological monthly mean distributions for each latitude band and pressure level are then determined by averaging over the 20 years. The test itself examines the difference between the model and climatological most probable temperature (MPT) for each month and latitude band, resulting in 132 points of comparison on each of 8 pressure surfaces. Figure 1 shows how the FVGCM and FVDAS MPTs differ from the NCEP/NCAR values on the 50 hPa and 5 hPa surfaces. At 50 hPa, both simulations do an excellent job of producing climatological temperatures; most differences range from 0-3K. At 5 hPa, both simulations are too cool, but the FVDAS is generally about 3K closer to the climatological temperatures than the FVGCM. Also, there is no apparent pattern to the FVDAS differences while the FVGCM's worst agreement proceeds from northern spring, through the tropics, to southern spring.

It is important to remember that this is not a comparison of model climatology with NCEP climatology. FVDAS temperatures were assimilated for the period July 1999-June 2000 while the FVGCM temperatures represent one year of a 35-year GCM simulation. Model temperatures cannot be judged as good or bad by this test; rather, this test is a general reality check to ensure that neither model year chosen deviates too far from observed climatology.

To summarize how each model level compares with observations, an area-weighted distribution of the differences between one year of model and climatological most probable temperatures are plotted in Figure 2. The FVDAS consistently produces MPTs in better agreement with climatology than the FVGCM. For the 8 pressure levels tested, the FVDAS is within 3K of climatology 82% of the time, while the FVGCM agrees within 3K 69% of the time. The results at 100 hPa are quite interesting because this is the only level with a bimodal distribution of differences. Contouring these differences as a function of latitude and time one easily sees the reason: both simulations have a bias toward low temperatures near the tropical tropopause, while poleward of 30° each model has excellent agreement with climatology.

2.1.2 Residual Circulation

Many GCMs cannot produce a quasi-biennial oscillation (QBO), so our tests of model circulation emphasize regions outside the tropics and subtropics, away from the secondary circulation set up by the QBO [*Jones and Pyle*, 1984]. We do wish to evaluate the mean behavior of tropical transport, so the test in Section 2.1.2.2 includes profile data from both QBO phases, and the test in Section 2.1.2.3 is a comparison that does not depend on the exact location of the subtropical boundary, which varies with the phase of the QBO.

2.1.2.1 Annual cycle of CH_4 in the extratropical middle and upper stratosphere

The transport characteristics of the middle and upper stratosphere are relevant to the ozone distribution even though photochemistry strongly influences ozone there. For example, NO_x family chemistry dominates O_3 loss in the sunlit middle stratosphere, but NO_x mixing ratios depend strongly on NO_y abundance, which is controlled largely by transport. The dynamics and composition of the upper stratosphere also affect ozone in the polar lower stratosphere because the strength of wave activity aloft determines descent rates and influences the fraction of upper

stratospheric air reaching the lower stratosphere by the end of winter [*Rosenfield and Schoeberl*, 2002].

We can evaluate model transport and mixing processes affecting the upper stratosphere through analysis of the CH_4 annual cycle. In the high latitudes it is a function of seasonally varying meridional transport, mixing, and descent. Photochemistry also matters in the summer upper stratosphere. The amplitude, phase, and variability of CH_4 annual cycle provide useful measures of the timing and strength of transport and the presence of photochemical processes. HALOE CH_4 data from 1992-1999 show large interannual variability equatorward of 44°, probably due to QBO influence, so we choose to study only the middle and high latitude ranges, 44°-56° and 72°-80° which show low interannual variability. CLAES CH_4 must be used for this test because HALOE does not sample the polar region in fall and winter.

This test compares several features of the CH_4 annual cycle at 800K (~10 hPa) and 1200K (~5 hPa). Probability distribution functions (pdfs) calculated from daily 1992 CLAES data within each latitude band are contoured together to show the amplitude, phase, and variability of the cycle. Figure 3 shows a one-year cycle of contoured pdfs of CH_4 from CLAES, the FVGCM, and the FVDAS at 1200K for 4 extratropical latitude bands. UARS yaw maneuvers result in 35-day gaps poleward of 34° in each hemisphere 5 times a year. Narrow distributions (yellow and red) indicate a homogeneous atmosphere, which can result from rapid photochemistry or mixing, while broad regions (purple, dark blue) indicate that transport from a photochemically different region dominates processes that homogenize the atmosphere. Four diagnostic quantities derived from data are described below; the scoring of GMI_{GCM} and GMI_{DAS} are summarized in Table 1.

Most Probable CH_4 mixing ratios. We compare the most probable value rather than the mean because it is not affected by the spatial distribution of the tracer and thus is more accurate

measure of a region's composition [*Sparling,* 2000]. In the real atmosphere, tracer mixing ratios depends on various transport processes as well as chemistry. In a model, particularly in the upper stratosphere, tracer mixing ratios also depend on the height of the model lid. For example, if the lid is near the stratopause, the model will not have a mesospheric source of low CH_4 to descend into the polar stratosphere in winter [*Ma and Waugh,* 2003]. While this test assesses whether the general balance between transport and chemistry is right, the results may also reflect the implementation of the model.

At each latitude band and height tested, the model receives a point for being within 25% of the observed value for the entire year, 0.5 point for being about 25% from the mean, and nothing for being more than 25% from the mean. In general, GMI_{GCM} mixing ratios are much lower than CLAES at 1200K, but show better agreement at 800K, especially in the NH. The GMI_{DAS} most probable values are usually quite close to the CLAES values and the total score for the GMI_{DAS} was much higher than the GMI_{GCM}.

Phase and amplitude of the annual cycle. This quantity reflects seasonal variations in the radiative forcing and wave driving. To evaluate these aspects of the circulation independently of the mean state, model results are first scaled by the ratio of the observed/model most probable annual values. After scaling, the amplitude is judged by whether the model follows the observed annual cycle and stays within 25% of the observed values for the year. For this it receives a full point, 0.5 point for not exceeding 25% of observed, and 0 for a cycle beyond $\pm 25\%$ of observed. The phase is judged by requiring the model to have a minimum and maximum within a month of observed, or, when appropriate, for correctly lacking a cycle (1 point). The model receives a half point for getting only the minimum or maximum right, and nothing for a phase that bears no resemblance the observations. Both simulations did extremely well.

Variability and its seasonal cycle. Long-range transport increases CH_4 variability while mixing and photochemistry decrease it. Variability has its own seasonal cycle independent of the seasonal cycle of the most probable value. For example, the midlatitude panels in Figure 3 show almost no variation in the most probable CH_4 value, yet the breadth of the distribution varies greatly between summer and winter. In the midlatitudes in both hemispheres, the pattern of CLAES CH_4 variability shows a minimum in summer and a broad maximum in fall and winter. This indicates that wave activity is strongest in fall and winter and weakest in summer. In the Antarctic, the pattern is similar but with smaller variability (weaker wave activity) near the vortex, especially at 800K in winter (not shown). In the Arctic, variability is greatest in winter and quite low in summer when photochemistry is fast. This semi-quantitative test looks at the breadth of the model distribution in each season and looks for agreement with the seasonal cycle in variability. The model values are scaled, as before, so that variability is judged independently of the annual mean.

The GMI_{GCM} overall shows a very good cycle of CH_4 variability at all latitudes and heights tested. Its cycles are good because there is always a minimum in variability in summer (i.e., wave-driven transport does not interfere with the reduction of variability by fast photochemistry) and a maximum in variability in the appropriate cool season. Its grades are less than perfect in the polar regions because the winter wave driving appears to be a little weak. This is consistent with the results of the MPV test, where low CH_4 at 1200K also pointed to weak wave driving. The GMI_{GCM} midlatitude cycle is in very good agreement with CLAES.

The GMI_{DAS} does not consistently show the right cycle of CH_4 variability. Like the GMI_{GCM}, wave driving in the polar regions is too weak, but it receives lower grades because of noisiness in summer, especially in the northern hemisphere. Like the GMI_{GCM}, the GMI_{DAS} does

best in the midlatitudes, with the exception of 1200K in the north. The minimum variability occurs too soon there (end of winter), and then it grows to be very large in spring and summer, looking nothing like the observations. The large summer variability seen in the northern midlatitudes in the GMI_{DAS} could be caused by excessive tropics-to-pole transport. This will be evaluated in Section 2.1.2.3.

2.1.2.2 Seasonal variations in lower stratospheric N_2O

Lower stratospheric profiles of the long-lived tracer N_2O reflect the balance between the diabatic circulation and meridional mixing there. N_2O measurements in the midlatitude and tropics, derived from 11 years of ER-2 N_2O data, are used to create seasonal mean profiles for this test [*Strahan et al.*, 1999]. This test was described in *Douglass et al.* [1999] (Test 2b) and the details of the grading can be found there. Figure 4 shows examples of the model/data agreement. The GMI_{GCM} scores higher than best simulation of GMI-1. Both simulations perform acceptably in the tropics and northern midlatitudes, and neither has excellent agreement in the southern hemisphere (SH). The GMI_{GCM} is closer to observations than the GMI_{DAS}.

2.1.2.3 Tropical isolation in the middle and upper stratosphere

This is a variation of the original 'Test 3' in *Douglass et al.* [1999], in which the bimodality of CLAES N_2O pdfs between 10°S and 45°N were used to assess a model's ability to produce a sufficient tropical/midlatitude barrier. The barrier is a steep gradient in potential vorticity that arises from midlatitude Rossby wave breaking [*Polvani et al.*, 1995]. The wave breaking, which causes mixing in the surf zone, cannot penetrate the tropics, resulting in distinctly different tracer mixing ratios in the two regions. Ascent of young air gives high tracer values in the tropics, while the older air of the middle and high latitudes has lower mixing ratios and a broad distribution. Seasonal variations in wave driving cause variations in the barrier strength, affecting

the distinctiveness of tropical and midlatitude air masses. The original test was conducted on 3 pressure surfaces from 31-7 hPa, while here we use 4 theta surfaces from 600-1200K. A wide latitude range is chosen so that QBO phase-dependent variations in the subtropical boundary will not affect the modality of the pdf. The phase-dependent secondary circulation set up by the QBO also causes significant interannual variability in constituent mixing ratios [*Jones and Pyle*, 1984]. With less than 2 years of CLAES N_2O data, we compare only the modality of the distribution and not the mixing ratios or the absolute separation of the peaks.

The models are graded every 200K between 600K and 1200K. The CLAES N_2O pdfs show isolation of the tropics at all levels and in all seasons compared. (Spring is excluded because the observed N_2O distribution was nonstationary.) A simulation is granted 1 point for producing 2 peaks separated by a minimum, even a weak one. A half point is given for a tropical peak with a long midlatitude tail instead of a clear minimum, and no points are given if a single, short-tailed (i.e., well mixed) peak is found. Figure 5 provides examples of the performance of these simulations at 800K. The GMI_{GCM} maintains a tropical/midlatitude separation in all 3 seasons (scoring 93%), while the GMI_{DAS} makes a clear separation only in winter (scoring 54%). The GMI_{GCM} performed equally well at all levels tested, while the GMI_{DAS} showed decreasing tropical isolation with increasing height.

2.1.3 Upper troposphere/Lower stratosphere Separation

This simple test is important because it gauges whether a model has the correct pathway of transport from the upper troposphere to lower stratosphere. Referred to as 'Test 5' in *Douglass et al.* [1999], it examines the phase lag of the CO_2 seasonal cycle across the tropopause at 60°N. Models must show at least a 2-month lag between the CO_2 seasonal cycle maximum on the

highest tropospheric and the lowest stratospheric levels. The presence of the lag indicates that air from the extratropical upper troposphere does not go directly up into the lower stratosphere, but takes a path to the stratosphere via the tropical tropopause [*Boering et al., 1*996; *Strahan et al.,* 1998]. This test is useful for identifying simulations with excessive convective transport. In this study, both simulations pass.

2.2 Specialized tests relevant to ozone simulations

2.2.1 Spatial and temporal coverage of PSC-producing temperatures

In regions where heterogeneous chemical reactions occur, the distribution of temperatures is more important than the mean. *Austin et al.* [2003] evaluated model polar temperatures by calculating the product of models' areal and temporal coverage of NAT- and ice-forming temperatures in each hemisphere and comparing with that quantity derived from observations. In this test we judge a model by whether it can produce a spatially realistic distribution of NAT-forming temperatures during the appropriate months. Since polar ozone loss depends on both low temperatures and sunlight, this test is designed to look specifically at the latitudes and months where PSC-forming temperatures are reached. The 'score' for this test is a description of the model's behavior compared to climatology.

This test uses 20 years of the NCEP/NCAR reanalysis to calculate a climatological temperature distribution for 3 latitude bands in each polar region (70^{o}-90^{o}, 60^{o}-70^{o}, and 50^{o}-60^{o}) at 6 pressure levels from 150-10 hPa. Since low temperature bias in the Antarctic stratosphere is a longstanding model issue [*Pawson et al.,* 2000], this test is useful not only to determine if a model makes enough PSCs deep in the Antarctic vortex, but to assess whether a model forms PSCs outside of the expected latitudes, heights, and seasons. This test labels a model's behavior as climatologically normal, somewhat warmer or colder than normal, or unrealistically warmer

or colder than normal.

We calculate the area-weighted fraction of each latitude band that is covered by temperatures at or below the PSC frost point (calculated with 4 ppmv H_2O) on each pressure surface in one month. Using 20 years of NCEP/NCAR analyses, the mean and standard deviation are calculated for the area-weighted fraction and compared with the same model-derived quantities. Model fractional coverage within 1.5 standard deviations (σ) of the climatological fraction is labeled 'normal'; fractional coverage between 1.5σ and 3σ above (below) the observed mean is considered 'colder (warmer) than normal', and 3σ above (below) the mean is 'much colder (warmer) than normal' and possibly unphysical.

Figure 6 characterizes polar temperature in the FVGCM and FVDAS from May to November. In fall, winter, and spring, both models produce climatologically normal temperatures from 50-90°S and 150-10 mb most of the time, but usually some part of the polar region is below normal during each month. The FVGCM tends to be a little too cold at 100 hPa and below, especially at latitudes near the edge and outside the vortex. Both models are too cold inside the vortex near 30 hPa in late winter, which could lead to greater ozone loss there. The information presented in this figure provides a context for the interpretation of ozone behavior in the *Considine et al.* [2003] and *Connell and Douglass* [2003].

The same analysis was performed for December through March in the northern hemisphere. The Arctic has higher mean temperatures and much larger interannual temperature variability than the Antarctic, such that being within 1.5σ of the mean can mean having some PSCs or none at all. Figure 7 characterizes the Arctic vortex temperature in the models. Much of the FVGCM vortex and edge region can be labeled 'normal'. Temperatures are on the warm side of the mean in early winter (no PSCs) while it is quite cold from 70-150 hPa in late winter. However, the

absolute effect on PSC areal coverage is small – for example, 'much colder than normal' at 100 hPa in March still means less than 10% coverage, compared to about 2% climatologically. The FVDAS temperatures are climatologically normal much of the time, with the exception that February and March are very much colder than average at lower levels. The FVDAS assimilated wind fields represent the period July 1, 1999-June 30, 2000, reflecting the unusually low lower stratospheric temperatures observed in March 2000 [*Newman et al.*, 2002].

The particular FVGCM year evaluated here, part of a 35-yr FVGCM integration, was selected because it was the most like the 'SOLVE' winter. It is interesting to note that this test rates the FVGCM winter as 'normal'. An average FVGCM winter from this multi-year integration would probably be rated 'warmer than normal'.

2.2.2 Lower stratospheric vortex behavior during breakdown

The huge ozone loss rates observed in the Antarctic vortex in spring are the result of a unique combination of dynamics and chemistry found nowhere else. A model's ability to realistically simulate vortex erosion and mixing processes during breakdown may be crucial to its credibility in predicting how declining halogens will affect the depth and the dispersion of Antarctic ozone loss. For example, if a model brings midlatitude air into the vortex in early spring, this intrusion of non-denitrified air will cause Cl-catalyzed loss processes to shut down prematurely. Ozone will not get realistically low inside the vortex, and less ozone-depleted air will be dispersed to lower latitudes. This test gauges vortex isolation and exchange between the vortex and midlatitudes in spring.

Methane measurements from HALOE and CLAES are both suitable for developing a mixing and isolation diagnostic, but the HALOE data provide a more complete picture. CLAES has good spatial coverage down to 80°S for a month at a time, but there are no measurements in

October and only 1 austral spring was sampled. CLAES CH_4 data also have large uncertainty at 450K where ozone loss rates are greatest [*Roche et al.*, 1996; *Considine et al.*, 2003]. While HALOE obtains only 15 profiles a day in a hemisphere, it has done so for a decade; low interannual variability in the southern hemisphere allows those measurements to be sensibly combined to derive a mean dynamical picture of vortex development and breakdown for the entire austral spring. The two CH_4 data sets can be compared on the 600K surface in September and November, where uncertainties are acceptable. In a prototype test, HALOE and CLAES CH_4 pdfs exhibited the same dynamical features, lending confidence to the use of HALOE data, which has far less spatial coverage in any single year.

The test examines the springtime evolution of the CH_4 pdfs of two latitude bands, 60-80°S and 40-60°S on the 450K and 600K surfaces. Pdfs are derived by binning 8 years of measurements from each latitude range for each month of spring. The 60-80°S range is almost strictly vortex air in early spring, retaining a small but isolated vortex core into November. The 40-60°S band is almost strictly midlatitude, or 'surf zone', air. The dynamics of vortex breakdown and the extent and direction of mixing between the vortex and surf zone are revealed by several features of the pdfs: the separation of the peaks, the depth of the minimum between the peaks, and changes in the means and most probable values of the peaks during spring.

The evolution of HALOE CH_4 pdfs, shown in the middle column of Figure 8, reveals the process of vortex breakdown in the Antarctic lower stratosphere. HALOE September data show two broad but distinct distributions with peaks separated by about 400 ppb; the lack of a deep minimum indicates a strong barrier to mixing has not yet developed. By October and November, the deep minimum indicates development of a very sharp boundary between the regions. Mixing within the vortex also increases during these months as indicated by the narrower vortex

distribution. The most probable value decreases, but as CH_4 decreases toward the pole, this may reflect the higher mean sampling latitude in October (69°S) compared to 64°S in September (67°S in November). The most probable value in the vortex declines more than 30 ppb between October and November, arguing strongly against any intrusion of surf zone air, which is typically 500 ppb higher in CH_4; even a narrow band of mixing at the vortex edge would result in an increase of the vortex mean. Diabatic heating is near zero in spring [*Rosenfield et al.*, 1994; *Rosenfield and Schoeberl*, 2001] and thus descent is not expected to be important. Notice that the high CH_4 peak of the 60-80°S November distribution is nearly identical to the 40-60°S peak, suggesting that air exiting the vortex becomes rapidly mixed into the surf zone. The development of this bimodal structure in the 60-80°S band and the endurance of the low CH_4 (vortex) peak indicate that 1) the vortex breaks down by erosion and, 2) breakdown appears to be a one-way process with no evidence of midlatitude air mixing into the vortex. The HALOE data at 600K, not shown, give a similar picture of breakdown and show an even stronger barrier to mixing.

This data set provides a clear picture of vortex breakdown and an excellent basis for model evaluation of this process. The left column of Figure 8 shows the breakdown of the GMI_{GCM} vortex. The evolution of the model breakdown differs in many ways from the observations. While the September distributions have an acceptable 400 ppb separation, that separation decreases in the following months, both distributions narrow (no long tails exist at any time), no deep and wide minimum forms between the peaks, and the most probable vortex value increases by nearly 200 ppb in stark contrast to the 100 ppb decrease in the observations. Vortex evolution in the GMI_{DAS} simulation is very similar. The separation of the GMI_{DAS} distributions is smaller to begin with and shrinks rapidly in spring. By November the GMI_{DAS} vortex and midlatitude distributions are strongly overlapping.

In both GMI CTM simulations, vortex evolution is clearly very different from reality. The decreasing separation of the peaks and the lack of a deep minimum between them indicate that a strong barrier to mixing never forms. The fact that the model vortex most probable value increases during spring suggests that midlatitude air is mixing into the vortex. The HALOE pdfs clearly show that the vortex breaks down by erosion rather than entrainment of midlatitude air, revealing a fundamental difference between modeled and observed vortex behavior. By November, the GMI_{DAS} CH_4 distributions indicate a nearly homogeneous region from 40-80°S, in great contrast to the HALOE pdfs that indicate the persistence of a small, well-isolated vortex. The large ozone losses observed each October in the Antarctic are possible only because of the strict isolation of chemically perturbed air inside the vortex. This unique requirement of both chemical processing and isolation cannot be met in a model that lacks a strong barrier to mixing at the vortex edge. Poor performance on this test suggests significant consequences for model vortex mixing ratios of Cl_y, ClO, and NO_x, and hence for ozone loss as well.

The likely reason for the increase in midlatitude HALOE CH_4 from 1370 to 1500 ppb during spring is wave-driven mixing in the surf zone, which brings high CH_4 poleward from low latitudes. At the same time, high latitude (vortex) CH_4 decreases, demonstrating the continued isolation of the vortex. The separation of the peaks of the HALOE vortex and midlatitude distributions increases during spring at 450K and 600K. In contrast to the HALOE analysis, both the GMI_{GCM} and GMI_{DAS} show mean CH_4 in the vortex increasing rapidly toward surf zone values in spring, suggesting a continuous exchange of air between the vortex and midlatitudes; the GMI_{DAS} has stronger exchange than the GMI_{GCM}. The models' mean values at 40-60°S are also increasing; further analysis showed this distribution merges with the 30-40°S distribution by November. The lack of vortex isolation in the simulations compromises their ability to sequester

chemically perturbed vortex air. The unphysical transport characteristics shown here may lead to significant consequences not only for the depth of ozone depletion, but for the timing of the dispersal of ozone-depleted air to midlatitudes during breakdown. In two experiments performed with FVGCM meteorological fields, one in another CTM and the other with CH_4 transported online in the FVGCM, the Antarctic vortex was well isolated. GMI-CTM details such as the horizontal resolution and the implementation of the *Lin and Rood* [1996] numerical transport scheme probably contribute to the transport characteristics displayed here.

Figure 9 compares model and HALOE Arctic pdfs on the 600K surface. There are no HALOE high latitude measurements in February. Although the Arctic vortex in March is much smaller than the Antarctic in September, the observations show a clear separation between the tiny vortex and the midlatitudes. The GMI_{GCM} manages to keep some separation between the air masses, but the GMI_{DAS} shows a completely homogenized region from 40-80°S by March; the observations indicate mixing is still incomplete in April. The GMI_{GCM} has broad distributions in February and March while the GMI_{DAS} distributions are much narrower, indicating stronger horizontal mixing in the GMI_{DAS}. In a typical Arctic winter with small O_3 losses, excessive mixing across the vortex edge will have little impact on the O_3 distribution since its horizontal gradients are fairly flat in the 500-600K range [*Strahan,* 2002]. However, should the Arctic stratosphere have a cooling trend in the 21st century with concomitantly larger wintertime O_3 losses, unrealistic vortex isolation such as shown here may invalidate model predictions.

3. Grading summary: model credibility

Most air enters the stratosphere through the tropical tropopause. Both simulations begin this journey reasonably well, with good agreement between modeled and observed N_2O profiles in

the tropical lower stratosphere. At 600K and above, the GMI_{DAS} is unable to maintain an isolated tropical air mass in summer and fall, unlike the GMI_{GCM} which maintains a distinct tropical air mass at all altitudes and seasons in tested. The weak tropical barrier in the GMI_{DAS} allows too much high N_2O into the midlatitudes, weakening meridional gradients there and resulting in lower stratospheric profiles with high N_2O (*e.g.*, Figure 4). The GMI_{DAS} tropical isolation gets worse with height. The excessive exchange between the tropics and midlatitudes in the GMI_{DAS} leads to problems with mid and high latitude tracer distributions and variability.

The test of CH_4 annual cycles in the southern extratropics shows insufficient wave driving in the austral fall. Wave driving brings high CH_4 from the low latitudes to the polar region. The inadequacy of the wave-driven transport is seen in the GMI_{GCM} Antarctic upper stratosphere in the form of too little CH_4, low variability, and almost no CH_4 increase in fall and winter compared to CLAES (Figure 3, 1200K 72-80°S). Similar CH_4 behavior is seen in an FVGCM simulation with online chemistry, suggesting the GMI-CTM implementation is not the cause of this disagreement with the observations. The GMI_{DAS} shows more of this transport occurring. When the GMI_{GCM} vortex forms in fall and descent begins in the upper stratosphere, CH_4 values lower than observed are trapped in the descending vortex. By late winter, CH_4 in the GMI_{GCM} Antarctic vortex is in close agreement with CLAES (600K-1200K), indicating some midlatitude air has mixed into the vortex during fall and winter; by spring, model CH_4 has become too high.

In Figure 8, showing the separation of vortex and midlatitude air, the vortex most probable value rises in the model while the decreasing in the observations. The GMI_{DAS} shows the same trend but starts with even higher CH_4 in the vortex. As previously discussed, these test results indicate that the lack of isolation allows too much high CH_4 air from middle latitudes into the vortex. However, because an online FVGCM experiment has demonstrated the model's ability to

produce nearly realistic vortex isolation year after year, we suspect that the lack of vortex isolation in the GMI simulations implicates the CTM implementation and/or horizontal resolution.

The overall result of the middle and high latitude residual circulation tests is that tracer transport in the offline GMI_{GCM} has greater fidelity throughout the stratosphere than the GMI_{DAS}. The GMI_{GCM} has greater realism in its northern hemisphere than its southern hemisphere and it performs best in the middle stratosphere. The southern hemisphere upper stratosphere is where the GMI_{GCM} has the worst comparison with observations and where the GMI_{DAS} scores considerably higher. The GMI_{DAS} performs best in the midlatitudes, north and south, but struggles with the southern hemisphere lower stratosphere. In the polar lower stratosphere, neither simulation is able develop an impermeable vortex edge. Table 2 summarizes residual circulation grading as a function of height and hemisphere, and rates barrier formation ability.

Temperature-dependent gas phase reactions are likely to be carried out at the right rates in both simulations. At 30 hPa and below, each model achieves climatologically normal temperatures at nearly all latitudes and seasons. Higher up, both models are biased slightly low, but near the stratopause both models are a few degrees too high. Overall, the FVDAS is always closer to NCEP 20-yr climatological temperatures than the FVGCM at all levels in the stratosphere.

Both simulations do a good job of producing near climatological temperatures for the production Antarctic PSCs, including realistic areal coverage as a function of month, latitude, and altitude. The region of FVGCM PSC-producing temperatures extends a little too far from the pole, particularly at 100 hPa and below. The FVDAS Anatarctic lower stratosphere warms a little later in spring than the climatological average. In the Arctic, the FVGCM was climatologically

near normal at 50 hPa and above, but cold in the late winter vortex at 70 hPa and below. The FVDAS was near normal in much of the Arctic, but much below normal in March outside the vortex near 70 hPa, reflecting the cold stratospheric spring of 2000.

Overall, the GMI_{GCM} does a better job of barrier formation, particularly in the tropics though only moderately so in the polar regions. This is consistent the model mean age of air comparison showing the GMI_{GCM} to have older air in the polar lower stratosphere than the GMI_{DAS} [*Considine et al.*, 2003]. The primary weakness in the GMI_{GCM} appears in the southern hemisphere spring and fall, the seasons when the greatest wave activity should occur [*Randel*, 1988]. The GMI_{DAS} does a better job there, suggesting that the insertion of observations in the DAS may improve GCM deficiencies in this region.

4. Conclusions

No model can faithfully represent all known atmospheric processes, but by understanding both the skills and the deficiencies of a model, one can determine its the best use. Transport and chemistry influence the distribution of ozone at all altitudes and latitudes of the stratosphere, requiring a model to perform well just about everywhere. To study the effect of changing chlorine levels on stratospheric ozone, a model requires additional testing in regions where chlorine plays a significant role in ozone loss (*i.e.*, the upper stratosphere and the polar lower stratosphere). This reflects the philosophy of evaluation used here.

These evaluations provide insight into the usefulness of offline chemistry and transport simulations using the FVGCM and FVDAS meteorological fields. The quality of these simulations is affected not only by the input meteorological fields, but by the offline model advection scheme, resolution of the CTM, and the implementation of various CTM components

(*e.g.*, the chemical mechanism). Differences in the time step of the chemical mechanism and the advection scheme will lead to interactions between these modules, especially for diurnally-varying species near the terminator. This leads to inherent differences in performance between offline and online chemistry. Experiments performed at 2^{o}x2.5^{o} horizontal resolution will not give the same results as a 4^{o}x5^{o} experiment, especially for meridional tracer gradients and for the CH_4 vortex mixing diagnostic. Experiments with online parameterized CH_4 chemistry in the FVGCM revealed that tracer transport is less diffusive and more realistic online than with the same meteorological fields in the offline model. Using the model diagnostics shown here, sensitivity of results to resolution and implementation choices can be tested, allowing the user to select simulations with the greatest fidelity to physical processes. Objective evaluation of model processes, using diagnostics such as these present here and in *Douglass et al.* [1999], provides a way to reduce uncertainty in model calculations.

Acknowledgments

We thank David Considine for insightful comments and suggestions. This work was supported by the NASA Atmospheric Chemistry, Modeling, and Analysis Program.

References

Austin, J. et al., Uncertainties and assessments of chemistry-climate models of the stratosphere, *Atmos. Chem Phys., 3*, 1-27, 2003.

Boering, K.A., S.C. Wofsy, B.C. Daube, J.R. Schneider, M. Loewenstein, J.R. Podolske, and T.J. Conway, Stratospheric mean ages and transport rates from observations of CO_2 and N_2O, *Science, 274*, 1340-1343, 1996.

Considine, D.B., P.S. Connell, D. Bergmann, D.A. Rotman, M.D. Fromm, J.D. Lumpe, and S.E. Strahan, Sensitivity of Antarctic ozone recovery predictions to GCM and DAS dynamics, *J. Geophys. Res., submitted*, 2003.

Danilin, M.Y., et al, Aviation fuel tracer simulation: model intercomparison and implications, *Geophys. Res. Lett., 25*, 3947-3950, 1998.

Douglass, A.R., M.J. Prather, T.M. Hall, S.E. Strahan, P.J. Rasch, L.C. Sparling, L. Coy, and J.M. Rodriguez, Choosing meteorological input for the global modeling initiative assessment of high-speed aircraft, *J. Geophys. Res., 104*, 27,545-27,564, 1999.

Connell, P.S. and A.R. Douglass, Chemical mechanism for the Global Modeling Initiative 3D chemistry and transport model, to be submitted, 2003.

Jones, R.L. and J.A. Pyle, Observations of CH_4 and N_2O by the Nimbus-7 SAMS – a comparison with in situ data and 2-dimensional numerical model calculations, *J. Geophys. Res., 89*, 5263-5279, 1984.

Kinnison, D.E. et al., The Global Modeling Initiative assessment model: Application to high-speed civil transport perturbation, *J. Geophys. Res., 106*, 1692-1712, 2001.

Lin, S.J. and R.B. Rood, Multidimensional flux form semi-Lagrangian transport schemes, *Mon. Weather Rev., 124*, 2046-2070, 1996.

Ma, J. and D.W. Waugh, Impact of model height on off-line stratospheric CTM simulations, *J. Geophys. Res., submitted*, 2003.

Newman, P.A., E.R. Nash, and J.E. Rosenfield, What controls the temperature of the Arctic stratosphere during the spring?, *J. Geophys. Res., 106*, 19,999-20,010, 2001.

Newman, P.A., et al., An overview of the SOLVE/THESEO 2000 campaign, *J. Geophys. Res., 107*, 8259, 2002.

Park, J.H., et al., Validation of halogen occultation experiment CH_4 measurements from the UARS, *J. Geophys. Res., 101*, 10,183-10,205, 1996.

Park, J.H., M.K.W. Ko, C.H. Jackman, R.A. Plumb, J.A. Kaye, and K.H. Sage, Models and measurements intercomparison II, *NASA/TM-1999-209554*, 1999.

Pawson, S., et al., The GCM-reality intercomparison project for SPARC (GRIPS): Scientific issues and initial results, *Bull. Am. Meteorol. Soc., 81*, 781-796, 2000.

Polvani, L.M., D.W. Waugh, and R.A. Plumb, On the subtropical edge of the stratospheric surf zone, *J. Atmos. Sci, 52*, 1288-1309, 1995.

Randel, W.J. Seasonal evolution of planetary waves in the southern hemisphere stratosphere and troposphere, *Q. J. R. Meteorol. Soc., 114*, 1385-1409, 1988.

Roche, A.E., et al., Validation of CH_4 and N_2O measurements by the CLAES instrument on the Upper Atmosphere Research Satellite, *J. Geophys. Res.*, 101, 9679-9710, 1996.

Rosenfield, J.E., P.A. Newman, and M.R. Schoeberl., Computations of diabatic descent in the stratospheric polar vortex, *J. Geophys. Res.*, 99, 16,677-16,689, 1994.

Rosenfield, J.E. and M.R. Schoeberl, On the origin of polar vortex air, *J. Geophys. Res.*, 106, 33,485-33,497, 2001.

Rotman, D.A., et al., Global Modeling Initiative assessment model: Model description,

integration, and testing of the transport shell, *J. Geophys. Res., 106*, 1669-1692, 2001.

Schoeberl, M.R., A.R. Douglass, Z.X. Zhu, and S. Pawson, A comparison of the lower stratospheric age spectra derived from a general circulation model and two data assimilation systems, *J. Geophys. Res., 108*, 4113, 2003.

Sparling, L.C., Statistical perspectives on stratospheric transport, *Rev. Geophys., 38*, 417-436, 2000.

Strahan, S.E., A.R. Douglass, J.E. Nielsen, and K.A. Boering, The CO_2 seasonal cycle as a tracer of transport, *J. Geophys. Res., 103*, 13,729-13,742, 1998.

Strahan, S.E., M. Loewenstein, and J.R. Podolske, Climatology and small-scale structure of lower stratospheric N_2O based on *in situ* observations, *J. Geophys. Res., 104*, 2195-2208, 1999.

Strahan, S.E., Influence of planetary wave transport on Arctic ozone as observed by Polar Ozone and Aerosol Measurement (POAM) III, *J. Geophys. Res., 107*, 4417, 2002.

World Meteorological Organization (WMO), Scientific assessment of ozone depletion: 2002, WMO 47, Geneva, Switzerland, 2002.

Figure Captions

Figure 1. The difference between the models' and the 20-yr NCEP climatological most probable temperatures on 2 surfaces. Most probable values are calculated monthly for 11 latitude bands. Contour intervals are 3K. At 50 hPa, both simulations are usually within 0-3K of the climatological value. The top panels show a cold bias in both simulations in the upper stratosphere (5 hPa). The FVDAS bias is smaller (mostly -3 to –6K), while the FVGCM bias is frequently –6 to -9K, with variable bias at high latitudes.

Figure 2. Summary of model temperature behavior with respect to the 20-yr NCEP climatology. Each histogram gives the area-weighted difference between model and climatological most probable temperatures (MPTs) on a given pressure surface for each latitude band and month of the year. (Each histogram contains 132 MPT comparisons.)

Figure 3. Comparison of CLAES CH_4 extratropical annual cycles with GMI_{GCM} and GMI_{DAS} for 4 latitude bands on the 1200K surface. The annual cycles are produced from contours of daily CH_4 pdfs. Yellow and red indicate a high probability of that mixing ratio, indicating a well-mixed distribution. Blue and purple, which represent low probability, are usually part of broad distributions. Broad distributions arise when long range transport dominates processes which reduce variability, namely, rapid photochemistry and mixing.

Figure 4. Comparison of seasonal mean N_2O profiles calculated from aircraft data with model profiles in the lower stratosphere over 3 latitude ranges and two seasons. Both models are consistently higher than the observations at and above 420K, with the GMI_{GCM} usually lying closer to the observations than the GMI_{DAS}.

Figure 5. Separation of tropical and midlatitude air masses illustrated by CLAES N_2O pdfs and model comparisons on the 800K surface. While both models maintain separation in winter at all

levels examined, only the GMI_{GCM} consistently keeps a clear separation in summer and fall at all levels.

Figure 6. Distribution of normal, below normal (1.5-3σ), and much below normal (>3σ) model temperatures in the Antarctic during the cold seasons. Both models have large areas of climatologically normal temperatures.

Figure 7. Same as Figure 6, except for the Arctic winter. Although the FVDAS is unusually cold in March, this accurately represents Arctic temperatures that year (2000). Overall, the FVGCM is warmer in the Arctic and makes fewer PSCs than the FVDAS, but because of the large variability there, both models are categorized as climatologically normal.

Figure 8. Evolution of CH_4 distributions on the 450K surface inside and outside the Antarctic vortex in spring. The central column, representing an 8-yr accumulation of HALOE observations in austral spring, demonstrates that the vortex air mass maintains its identity while gradually eroding. The GMI_{GCM} simulation (left column) maintains some separation through the spring, but large overlap between the distributions indicates exchange between the regions – in contrast to the observations. The GMI_{DAS} simulation (right column) does a worse job of maintaining separation, and by November the vortex and midlatitudes are nearly identical (*i.e.*, well mixed).

Figure 9. Evolution of CH_4 distributions on the 600K surface inside and outside the Arctic vortex in spring. The GMI_{GCM} distributions (left column) show good separation in February and March, with near total mixing by April. HALOE data, accumulated over 8 years (center column), show a small, distinct vortex in April. The GMI_{DAS} simulation (right column) cannot keep the regions distinct even in February. A substantial vortex existed all winter in 2000 [*Newman et al.*, 2002], but the GMI_{DAS} vortex and midlatitudes are indistinct in this simulation.

Table 1. Residual Circulation Test Results (from Section 2.1.2)

Height	Antarctic (72-80°S)	GCM	DAS
800K	MPV CH_4	0.5	0.5
800K	Phase and Amplitude	2.0/2	1.5/2
800K	Variability	0.75	0.5
1200K	MPV CH_4	0.0	0.5
1200K	Phase and Amplitude	2.0/2	1.5/2
1200K	Variability	0.75	0.5
	TOTAL	**75%**	**63%**
	Midlatitudes (35-56°S)	**GCM**	**DAS**
380-500K	Descent/Horizontal Mixing	0.67	0.33
800K	MPV CH_4	0.5	1.0
800K	Phase and Amplitude	2.0/2	2.0/2
800K	Variability	1.0	1.0
1200K	MPV CH_4	0.0	1.0
1200K	Phase and Amplitude	1.0/2	2.0/2
1200K	Variability	1.0	0.75
	TOTAL	**69%**	**90%**
	Tropics (10°S-20°N)	**GCM**	**DAS**
380-480K	Ascent/Horizontal Mixing	0.80	0.85
600K	Isolation	2.5/3	2.5/3
800K	Isolation	3.0/3	2.0/3
1000K	Isolation	2.5/3	1.5/3
1200K	Isolation	3.0/3	0.5/3
	TOTAL	**89%**	**60%**
	Midlatitudes (35-56°N)	**GCM**	**DAS**
380-500K	Descent/Horizontal Mixing	0.83	0.73
800K	MPV CH_4	1.0	1.0
800K	Phase and Amplitude	2.0/2	2.0/2
800K	Variability	0.85	0.75
1200K	MPV CH_4	0.5	1.0
1200K	Phase and Amplitude	2.0/2	2.0/2
1200K	Variability	1.0	0.25
	TOTAL	**91%**	**86%**
	Arctic (72-80°N)	**GCM**	**DAS**
800K	MPV CH_4	1.0	0.5
800K	Phase and Amplitude	2.0/2	1.5/2
800K	Variability	0.85	0.5
1200K	MPV CH_4	0.5	1.0
1200K	Phase and Amplitude	2.0/2	1.5/2
1200K	Variability	0.75	0.3
	TOTAL	**89%**	**66%**

Table 2. Summary of Transport Performance

Comparison of Simulations by Height and Hemisphere

	SH - FVGCM	SH – FVDAS	NH - FVGCM	NH - FVDAS
Upper Strat	59%	78%	84%	76%
Middle Strat	84%	81%	96%	78%
Lower Strat	67%	33%	83%	73%

Quality of Mixing Barriers (Very good, good, fair, poor)

	FVGCM	FVDAS
Antarctic Vortex	Fair	Poor
Arctic Vortex	Fair	Poor
Tropics (600-800K)	Very good	Fair
Tropics (1000-1200K)	Very good	Poor
Tropopause – 60°N	Very good	Very good

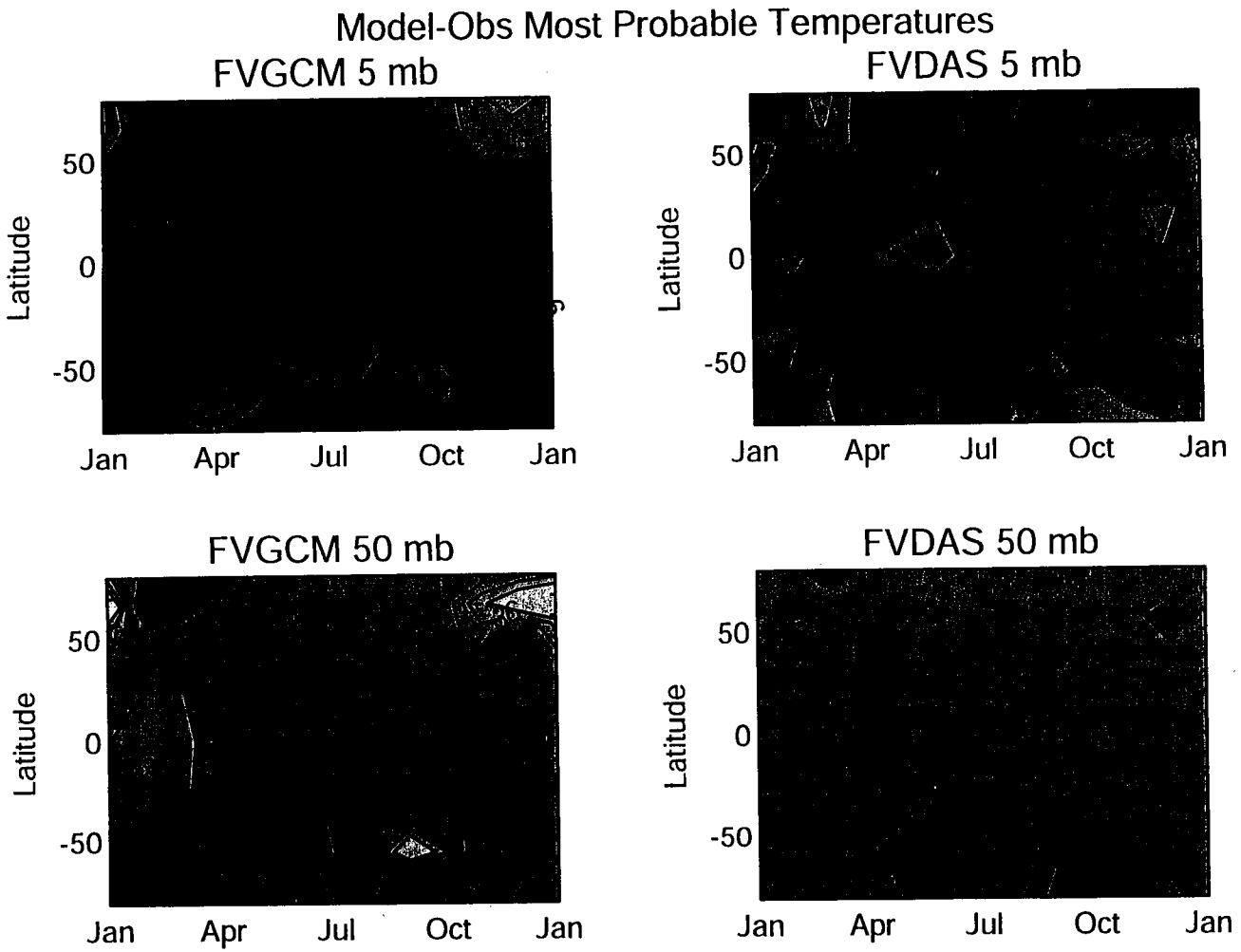
Model-Obs Most Probable Temperatures
FVGCM 5 mb
FVDAS 5 mb
FVGCM 50 mb
FVDAS 50 mb
Latitude
50
0
-50
Jan
Apr
Jul
Oct
Jan

Figure 1

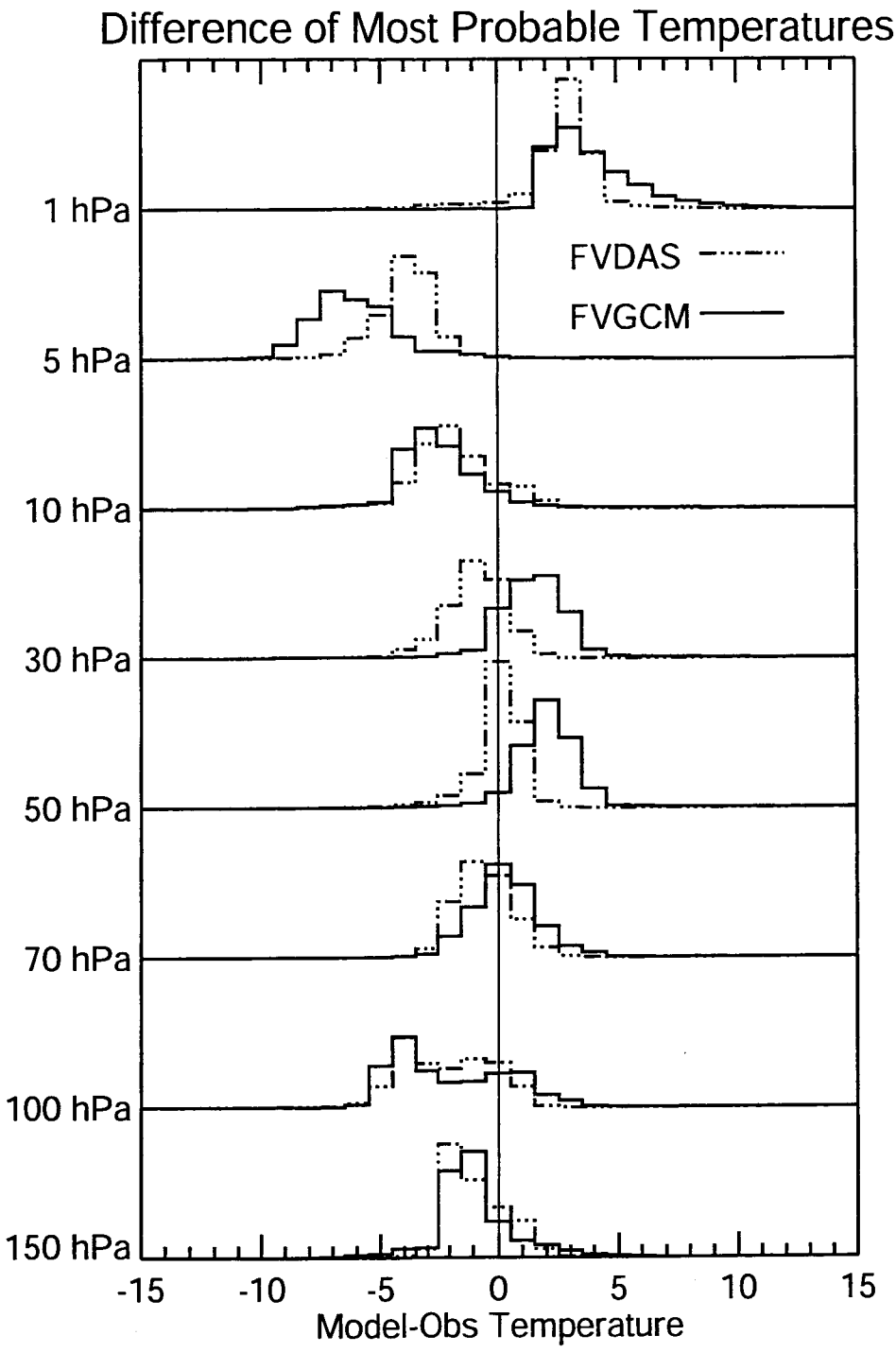
Difference of Most Probable Temperatures
1 hPa
5 hPa
10 hPa
30 hPa
50 hPa
70 hPa
100 hPa
150 hPa
FVDAS
FVGCM
-15
-10
-5
0
5
10
15
Model-Obs Temperature

Figure 2

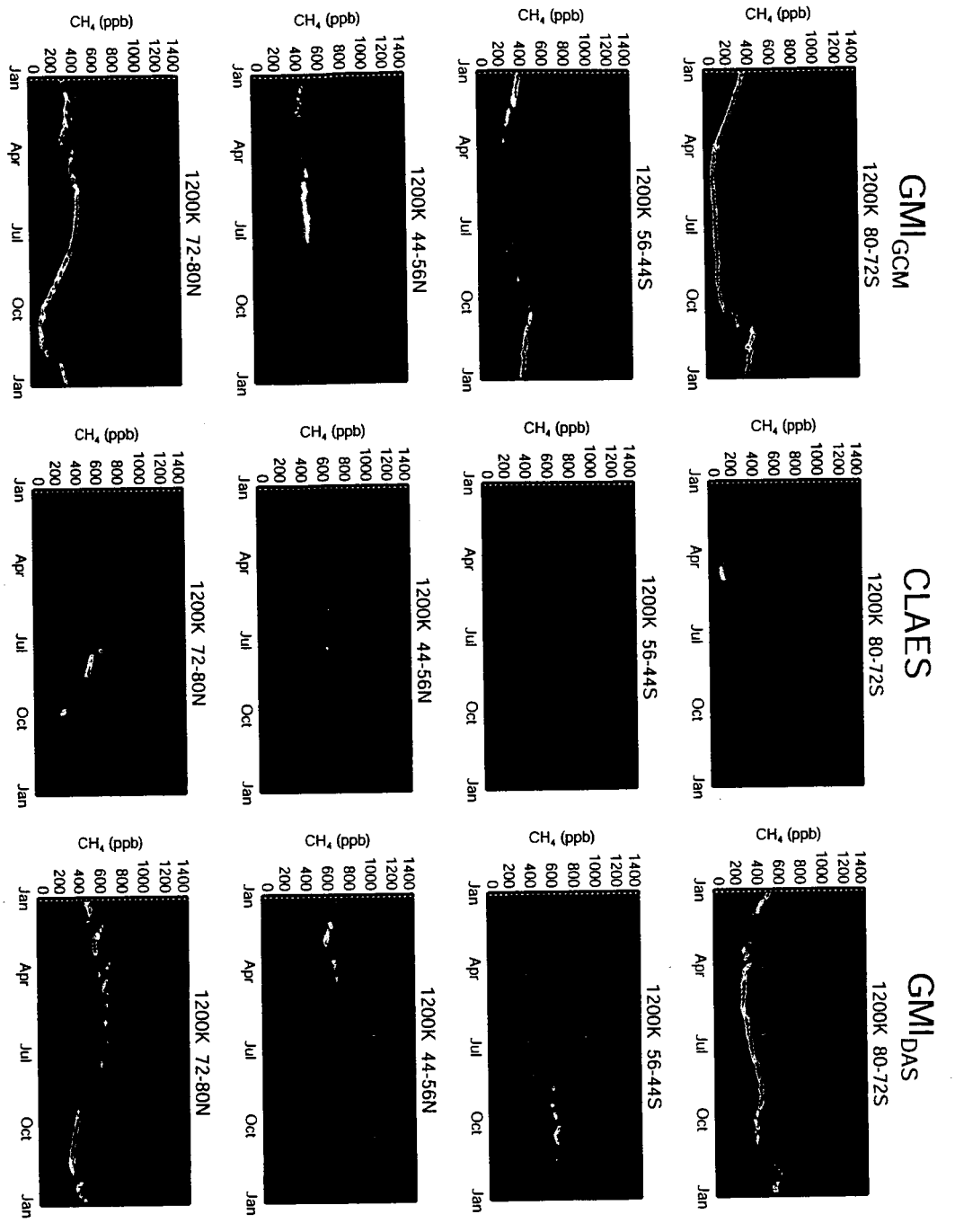
GMI_GCM
CLAES
GMI_DAS
1200K 80-72S
1200K 56-44S
1200K 44-56N
1200K 72-80N
CH4 (ppb)
0
200
400
600
800
1000
1200
1400
Jan
Apr
Jul
Oct
Jan

Figure 3

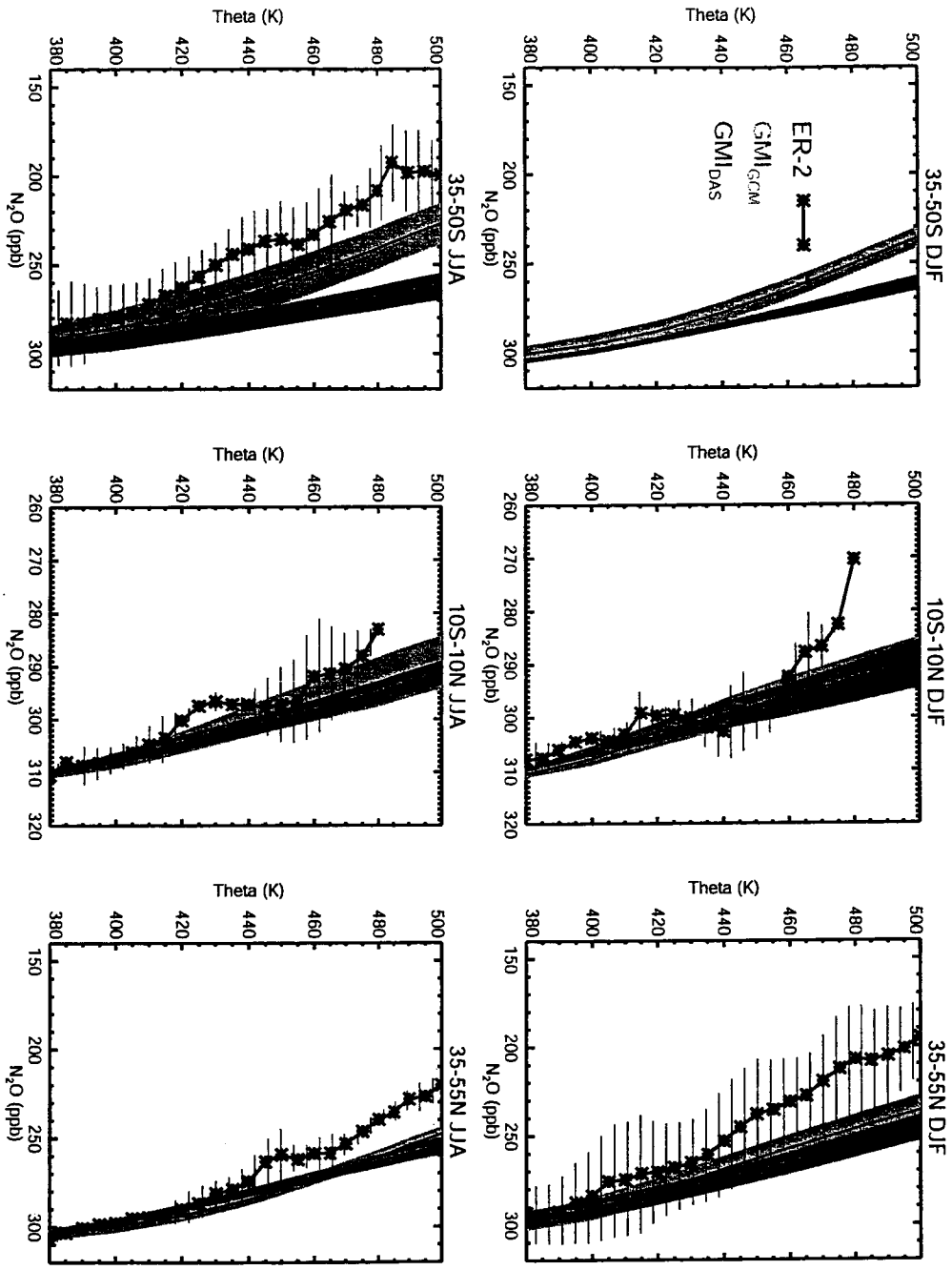
35-50S DJF
35-50S JJA
10S-10N DJF
10S-10N JJA
35-55N DJF
35-55N JJA
Theta (K)
N2O (ppb)
ER-2
GMIGCM
GMIDAS

Figure 4

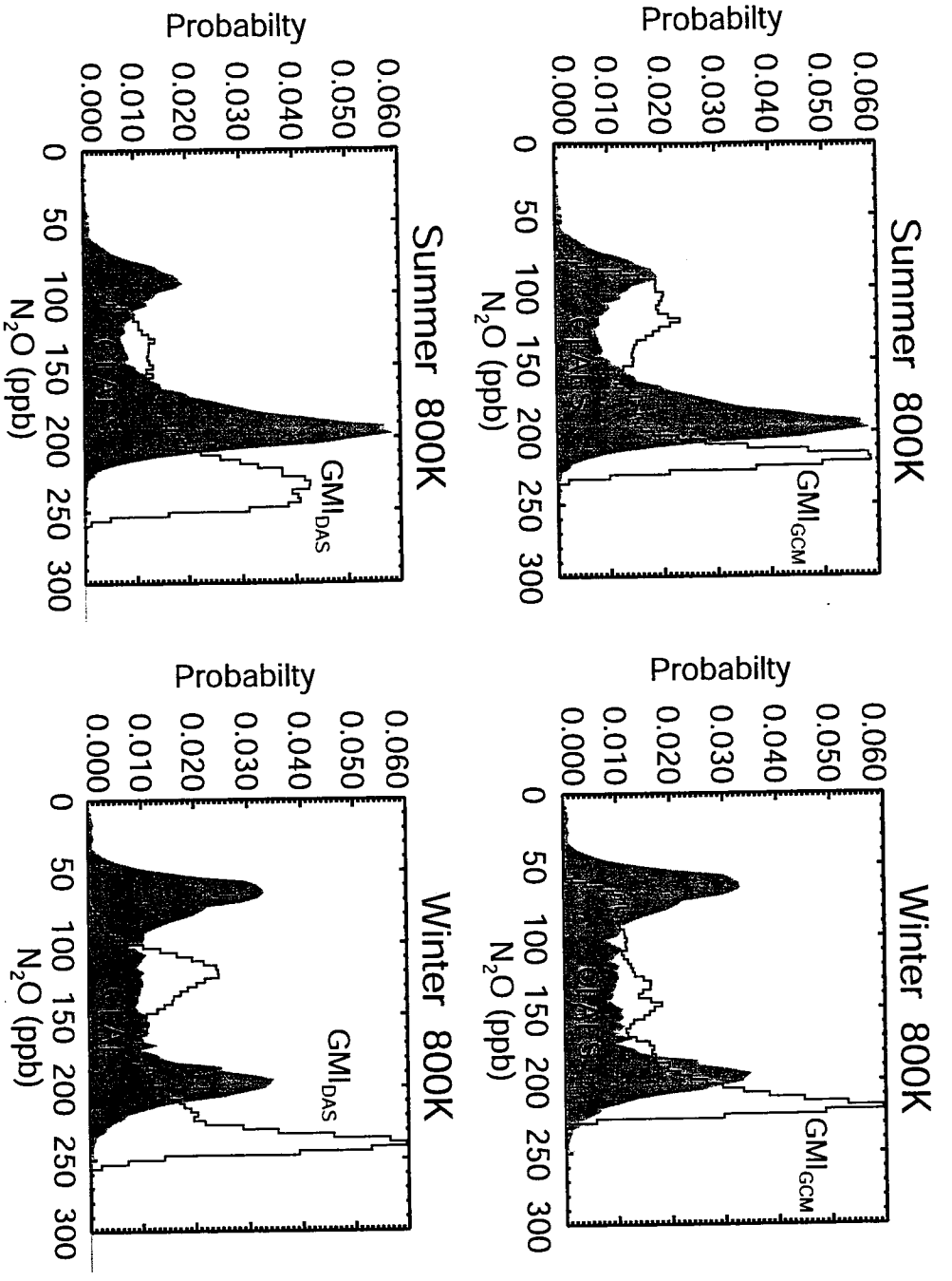
Summer 800K
Winter 800K
Summer 800K
Winter 800K
Probabilty
Probabilty
Probabilty
Probabilty
0.060
0.050
0.040
0.030
0.020
0.010
0.000
0 50 100 150 200 250 300
N2O (ppb)
GMIGCM
GMIDAS

Figure 5

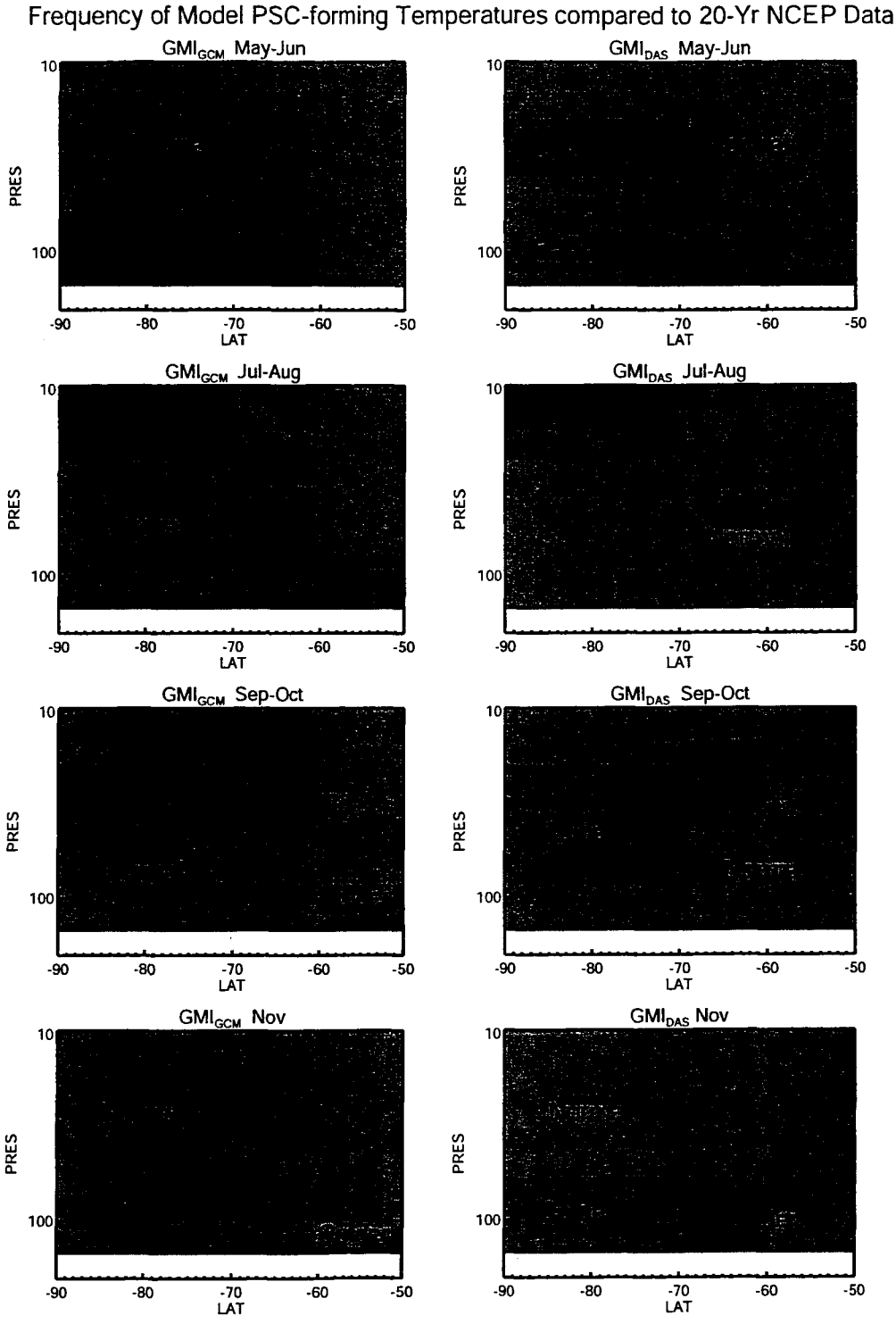
Frequency of Model PSC-forming Temperatures compared to 20-Yr NCEP Data
GMI$_{GCM}$ May-Jun
GMI$_{DAS}$ May-Jun
GMI$_{GCM}$ Jul-Aug
GMI$_{DAS}$ Jul-Aug
GMI$_{GCM}$ Sep-Oct
GMI$_{DAS}$ Sep-Oct
GMI$_{GCM}$ Nov
GMI$_{DAS}$ Nov
PRES
10
100
-90
-80
-70
-60
-50
LAT

Figure 6

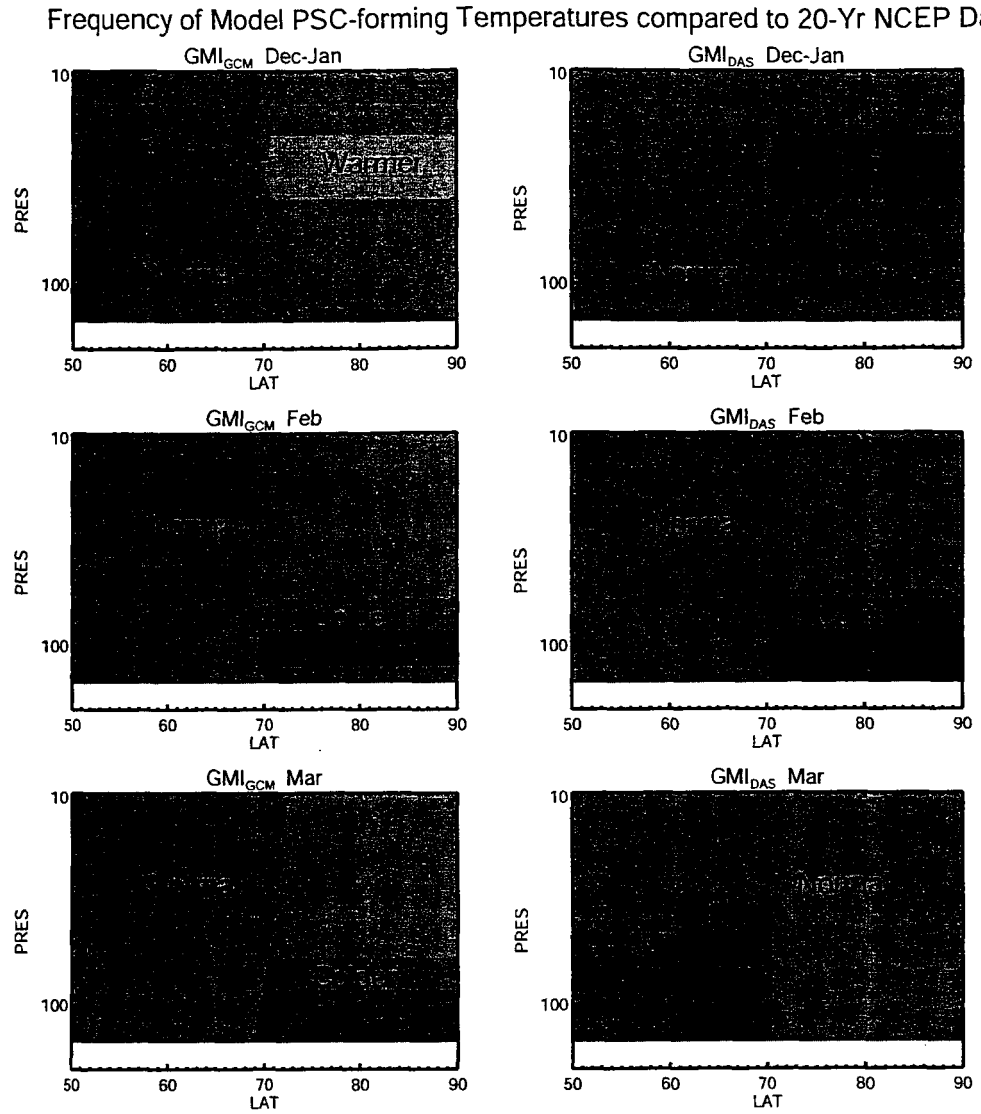
Frequency of Model PSC-forming Temperatures compared to 20-Yr NCEP Data
GMI$_{GCM}$ Dec-Jan
GMI$_{DAS}$ Dec-Jan
GMI$_{GCM}$ Feb
GMI$_{DAS}$ Feb
GMI$_{GCM}$ Mar
GMI$_{DAS}$ Mar
Warmer
PRES
10
100
50
60
70
80
90
LAT

Figure 7

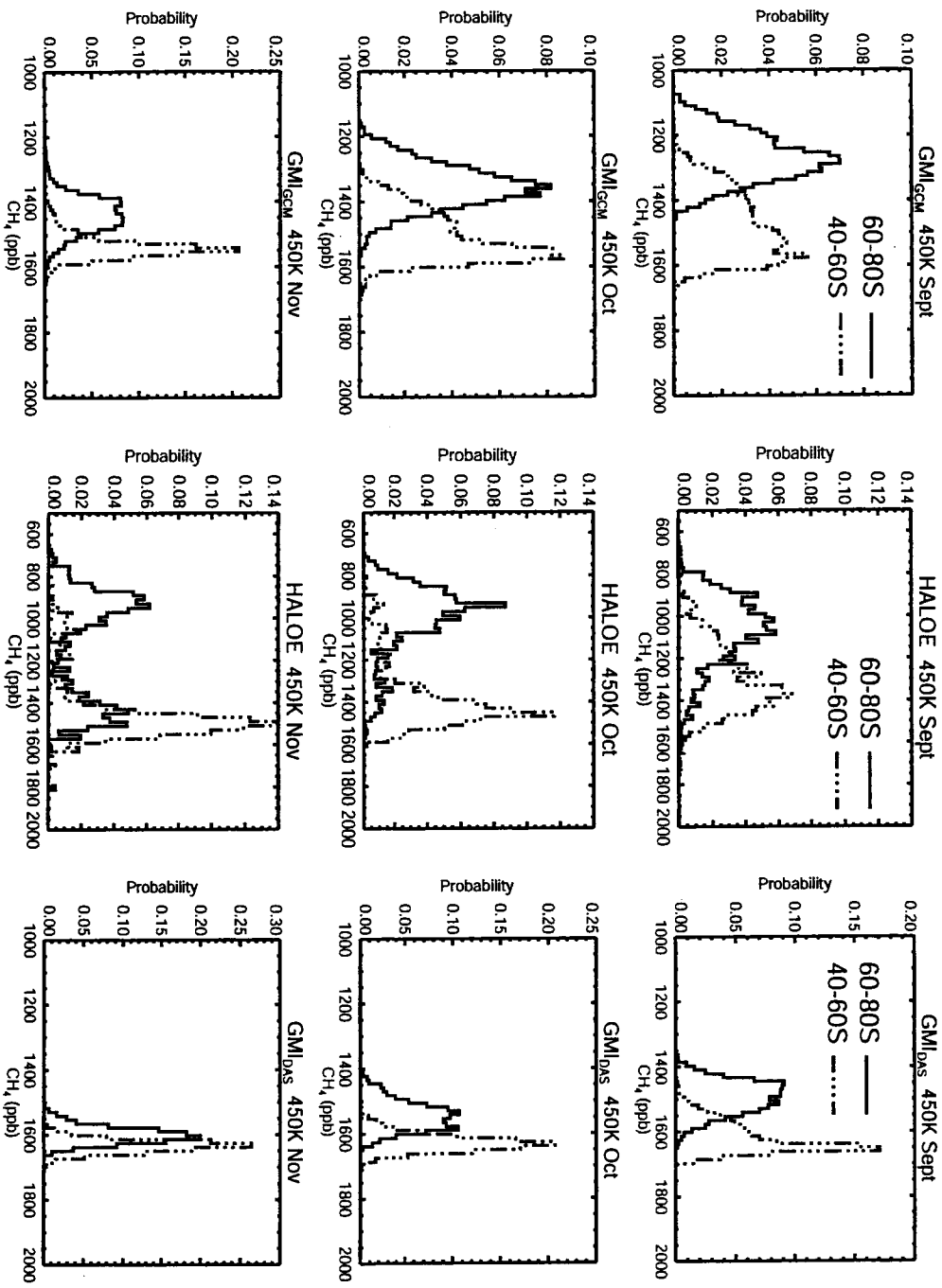
GMI_GCM 450K Sept
60-80S
40-60S
GMI_GCM 450K Oct
GMI_GCM 450K Nov
HALOE 450K Sept
HALOE 450K Oct
HALOE 450K Nov
GMI_DAS 450K Sept
GMI_DAS 450K Oct
GMI_DAS 450K Nov
Probability
CH4 (ppb)

Figure 8

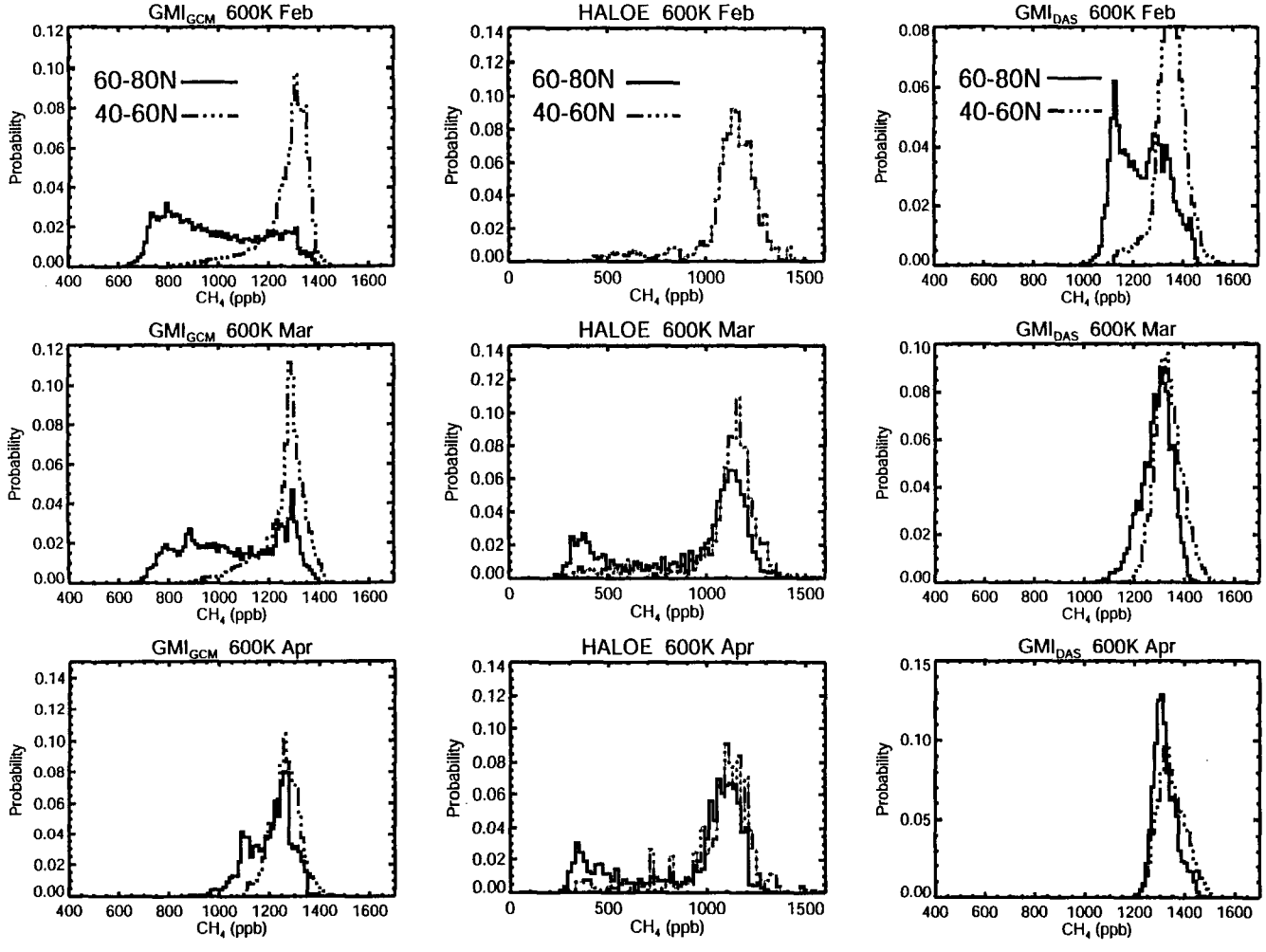
GMI$_{GCM}$ 600K Feb
HALOE 600K Feb
GMI$_{DAS}$ 600K Feb
GMI$_{GCM}$ 600K Mar
HALOE 600K Mar
GMI$_{DAS}$ 600K Mar
GMI$_{GCM}$ 600K Apr
HALOE 600K Apr
GMI$_{DAS}$ 600K Apr
60-80N
40-60N
Probability
CH$_4$ (ppb)

Figure 9